PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-350669
(43)Date of publication of application : 21.12.2001

(51)Int.Cl. G06F 12/08
G06F 3/06
G11B 20/10

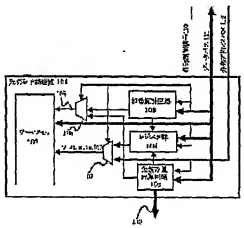
(21)Application number : 2000-175556
(22)Date of filing : 07.06.2000

(71)Applicant : HITACHI LTD
(72)Inventor : IGUCHI SHINYA
TSUNODA MOTOYASU
HONDA KYOSHI
ICHIKAWA MASATOSHI
TAKAYASU ATSUSHI
NISHIKAWA MANABU

(54) PRE-READ PREDICTING DEVICE

(57)Abstract:
PROBLEM TO BE SOLVED: To provide a pre-read predicting device controlling a pre-read quantity and obtaining an access characteristic for every region on memory media for improving the utilization efficiency of a cache memory in a disk device with the cache memory mounted.
SOLUTION: This pre-read predicting device comprises a weighted statistical circuit, a register group, a pre-read calculating circuit, and a work memory. The command data sent via a data bus are retained in the work memory, and the command history retained in the work memory is statistically processed by the weighted statistical circuit in consideration of the order of the command history. The statistical result is compared with the sent command by a pre-read quantity calculating circuit to calculate a pre-read quantity. In statistical processing, an evaluated value for every region of the memory media is retained in the work memory as reference for a disk control device in the data substituting work of the cache memory.

FIL



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁（ＪＰ）　(12)公開特許公報(A)　(11)特許出願公開番号
特開2001－350669
（P2001－350669A）
(43)公開日　平成13年12月21日(2001.12.21)

(51)Int.Cl.⁷	識別記号	FI	テーマコード（参考）
G06F 12/08	505	G06F 12/08	505C 5B005
	543		543B 5B065
	557		557 5D044
3/06	302	3/06	302A
G11B 20/10		G11B 20/10	A

審査請求　未請求　請求項の数5　OL　（全 26 頁）

(21)出願番号　特願2000－175556(P2000－175556)

(22)出願日　平成12年6月7日(2000.6.7)

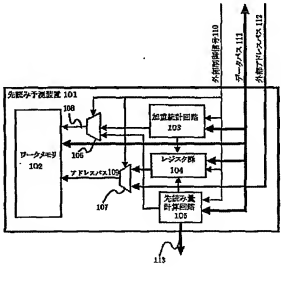
(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目6番地
(72)発明者　井口　慎也
神奈川県川崎市麻生区王禅寺1099番地　株式会社日立製作所システム開発研究所内
(72)発明者　角田　元泰
神奈川県川崎市麻生区王禅寺1099番地　株式会社日立製作所システム開発研究所内
(74)代理人　100075096
弁理士　作田　康夫

最終頁に続く

(54)【発明の名称】　先読み予測装置

(57)【要約】
【課題】キャッシュメモリを搭載したディスク装置において、キャッシュメモリの利用効率を改善するため、先読み量の制御と、記録媒体上の領域ごとのアクセス特性を求める先読み予測装置を提供する。
【解決手段】加重統計回路、レジスタ群、先読み量計算回路および、ワークメモリから構成された先読み予測装置を構成し、データバスを介して送られてきたコマンドデータをワークメモリへ保持すると共に、ワークメモリへ保持されたコマンド履歴を加重統計回路でコマンド履歴の順序を考慮した統計処理を行い、その統計結果と送られてきたコマンドを先読み量計算回路で比較し、先読み量を算出する。また、統計処理を行う際、ディスク制御装置がキャッシュメモリのデータの置換などの作業を行う際参照できるように、記録媒体の領域ごとの評価値をワークメモリへ保持しておく。

図1



【特許請求の範囲】
【請求項１】現在のコマンドと以前のコマンドの履歴から最適な先読み量を算出する先読み予測装置において、コマンドの履歴と、コマンド履歴内の各コマンドの順序、コマンドがアクセスしたロジカルブロックアドレス及びセクタ数によって重み付けして統計処理を行った結果を保持しておくワークメモリと、コマンド履歴のアクセス順序から重み付け統計処理を行う際に用いる評価値を求めるための評価値変換テーブルと、コマンドの履歴をワークメモリから取り出し、評価値変換テーブルを用いて評価値を求め、統計データと統計データの範囲の広がりを計算し、ワークメモリへ統計データを記録する加重統計計算回路と、統計データの範囲の広がりを先読み量の比率に変換する比較変換テーブルと、ワークメモリ内の統計データと現在のコマンドを比較し対応する統計データの先読み量に関するデータを取り出し、比較変換テーブルを用いて算出した先読み比率と登をとることで先読み量を算出する先読み計算回路からなることを特徴とする先読み予測装置。
【請求項２】請求項１記載の先読み予測装置を内蔵し、ホストとのインタフェースを制御するインタフェース制御回路、複数のアドレスバスを制御するアドレスマルチプレクサ、データバスを制御するデータマルチプレクサ、記録媒体を制御するディスクフォーマッタ、エラー訂正を行うECC、各モジュールを制御するシーケンサから構成され、ホストからコマンドを受け取ると直接先読み予測装置を起動して先読み量を計算し、ディスクフォーマッタに最大先読み量を設定することで記録媒体からの先読みを制限することを特徴とするディスク制御装置。
【請求項３】請求項１記載の先読み予測装置と、ディスク制御装置、キャッシュメモリ、ROM、CPU、記録媒体を制御するサーボ系制御回路から構成され、ディスク制御装置がホストからコマンドを受け取ると、CPUに対してディスク制御装置が割り込みを発生させることでコマンドの入力を知らせ、CPUが先読み予測装置を制御して先読み量を算出し、ディスク制御装置に対して先読み量を設定することを特徴とする記録装置。
【請求項４】請求項２記載のディスク制御装置を内蔵した記録装置において、CPUがキャッシュメモリ内のデータをセグメント単位で扱い、個々のセグメント情報を記録したセグメント管理テーブルによってセグメントを管理するキャッシュメモリの制御方式において、先読み予測装置が計算した評価値をCPUが読み出し、その評価値をセグメント管理テーブルの各セグメントに関する評価値情報と比較し、最も評価値が低いセグメント情報を破棄することで、先読みを行うデータの格納領域をキャッシュメモリ上に確保することを特徴とする記録装置。
【請求項５】請求項４記載の記録装置において、ホストの発行したコマンドが、キャッシュメモリ内のデータの一部にヒットした場合（ハーフヒット）、先読み予測装置が計算したアクセス評価値を調べ、評価値が高ければハーフヒットしたセグメントに不足分のデータを記録媒体から読み出すことで新しいセグメントを生成することなくハーフヒットしたセグメントのセグメントサイズを拡大することで、評価値の高いセグメントを残すように制御を行うことを特徴とする記録装置。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】本発明は、記憶装置及びそれを搭載する情報機器に関係し、特に記録装置からのデータの先読みの制御方式に関する。
【０００２】
【従来の技術】磁気ディスクを用いた記憶装置において、磁気ディスクのシーク動作などを考慮すると、アクセス速度はそれを使用するホストのデータ転送速度と比較してかなり低速である。このため、従来から磁気ディスクを用いた記憶装置にはこれらの速度差を吸収するためにキャッシュメモリが搭載されてきた。これを用いて、一度磁気ディスクから読み出されたデータをキャッシュメモリへ保存し、さらに、ホストが一度要求したデータの論理ブロックアドレス（以下、LBA）に対応するデータを磁気ディスクよりあらかじめ読み出す先読みと呼ばれる機能を用いることで、ホストへのデータ転送速度を向上させてきた。従来技術では、特開平6－119244に記載されているように、記録装置の物理領域ごとのヒット判定の結果を統計処理し、記録装置からの読み出しを制御する方式、特開平8－134634に記載されているように、ホストからのアクセスパターンがランダムアクセスか、シーケンシャルアクセスかを識別して、記録装置からの読み出しを行う方式、そして、米国特許5765213に記載されているように、キャッシュメモリ内のデータの格納状態から次に先読みを行うべき位置と先読み量を計算する方式がある。
【０００３】
【発明が解決しようとする課題】上記従来技術では、先読みを行う記録媒体上の各領域において、過去平均どれだけのデータが一度にアクセスされたかという情報を持たない。そのため、ヒット率は高いがホストが少量のデータしか要求してこないような領域に対してのアクセスが行われた場合、その領域に対して先読みが起動されると、必要以上にシーケンシャルに多量のデータをキャッシュメモリへ読み込んでしまう。このため、キャッシュメモリ上の他のデータを上書きしてしまうので、全体的に見てキャッシュメモリのヒット率が低下する。
【０００４】また、上記従来技術では、ホストからのシーケンシャルアクセスが一定期間続いた後、ランダムアクセス傾向が強い領域のデータのアクセスを行った場合、シーケンシャルに多量にキャッシュメモリに読み込んでしまい、他のデータを上書きしてしまうので、キャ

ッシュメモリのヒット率が低下する。

【０００５】また、上記従来技術では、先読制御のためのデータをキャッシュメモリの状態を処理することで得ていたため、キャッシュメモリの容量に依存する。このため、キャッシュメモリの容量が少ないと、先読み制御のためのデータを十分に得ることが出来ず、先読み制御の精度が低下する。

【０００６】本発明の目的は、キャッシュメモリ内のデータの不要な上書きを抑制し、キャッシュメモリの利用効率を改善することである。

【０００７】また、本発明の別の目的は、キャッシュメモリ内のデータのヒット率を向上させことである。

【０００８】さらに、本発明の別の目的は、キャッシュメモリの容量とその制御方式の影響を受けずに先読みの制御を行うことである。

【０００９】

【課題を解決するための手段】上記目的を達成するために、ホストからのコマンドの履歴とそのコマンド履歴から計算された統計結果を記録するワークメモリと、コマンド履歴をワークメモリから読み出し、コマンド履歴の順序により重み付けした統計処理を行い、記録装置の各領域に対する評価値と平均アクセス量を求めワークメモリへ統計結果を記録する加重統計処理回路と、その統計結果を読み出してホストからのコマンドと比較して、先読み量とホストが要求している記録媒体上の領域の評価値を計算する先読み量計算回路と、加重統計回路と先読み量計算回路で使用する値を一時的に保持するレジスタ群から構成された先読み予測装置を構成する。

【００１０】また、上記他の目的を達成するために、先読み予測装置を搭載し、ホストからのリード要求に対して、ホストが要求している記録媒体上の領域の評価値を求めると同時に、先読み予測装置がディスクフォーマッタに対して先読み制御のための指示を出し、先読み量を制限する機能を有するディスク制御装置を構成し、このディスク制御装置からCPUが評価値を取り出し、キャッシュメモリ内に保持されているデータの置換に用いることでキャッシュメモリを制御する記録装置を構成する。

【００１１】

【発明の実施の形態】図１は、本発明を用いて構成された先読み予測装置の構成図の一例である。先読み予測装置101は、統計データとコマンド履歴を保持するワークメモリ102と、ホストからのコマンド履歴から重み付け統計処理を行いアクセス特性を求める加重統計回路103と、統計結果から先読み量の算出と評価値を求める先読み計算回路105と、加重統計回路103と先読み量計算回路105が使用するレジスタをまとめたレジスタ群104から構成されている。尚、図２に示すようにワークメモリ102を先読み予測装置201に内蔵せず、外部メモリを使用してもよい。信号マルチプレクサ106は加重統計回路103と先読み量計算回路105の制御信号を外部制御信号110によって切り替えてワークメモリ制御信号108を作り出す。アドレスマルチプレクサ107は加重統計回路103と先読み量計算回路105、外部アドレスバス112を外部制御信号110によって切り替えてワークメモリ102へのアドレスバス109に接続する。データバス111は、各回路とワークメモリ102および外部回路のデータバスとの接続に使用されデータの送受を行うために使用される。

【００１２】本発明で行う統計処理の概念について説明する。図３は、本発明で行う統計処理の説明図である。

【００１３】図示のグラフは、縦軸を評価値301、横軸をLBA302としている。また、このとき、ホストからのコマンド履歴がコマンド順序303の順番で古い方から305a～305eの順番に保持されているとする。

【００１４】コマンド305aのLBAがA'でアクセスセクタ数がA'セクタとする。この時、グラフ上では、コマンド履歴の番号に応じた評価値分だけ、位置Aから始まる領域A'の評価値が上昇する。次に、コマンド305bに関しても同様の処理を行う。ただし、コマンド305aの領域と一部重なるため、重なっている両端でのLBAを求め記録する必要がある。コマンド305aとコマンド305bの重なりの先頭部分はコマンド305bのLBAであるBが対応するが、コマンド305aとコマンド305bの重なりの終端部分ついては、対応するLBAが存在しないため、新たにLBAを計算し、この計算した値を新たな統計データとして保持しておく。尚、図３ではこの新規LBAをaとしている。同様な手順でコマンド305eまでの重み付け統計処理を行った結果が図３の統計結果304である。

【００１５】図４は、実際の統計処理を行う場合におけるワークメモリ102の状態と各領域でのデータの形式、および計算式を示している。ワークメモリ102は、コマンド領域401と統計データ領域402および作業領域403に分けられる。コマンド領域401では、データ形式404に示すように、ホストからのコマンド履歴のLBAとアクセスセクタ数を古いものから順に保持している。本実施例では、数百個のコマンド履歴を保持できるとする。ただし、コマンド領域は有限であるので、コマンド領域にコマンド履歴が収まりきらなくなった場合もっとも古いデータを削除し新しいデータを書き込む。このためコマンド領域はリングバッファになっている。

【００１６】統計データ領域402はデータ形式405で示す形式で統計データを記録している。本実施例では、コマンド領域の約4倍程度の大きさを持つとする。評価値は数式(1)で示す計算式で計算される。数式(1)において、コマンド履歴と評価値の対応づけを行うためのテーブルについては後述する。LBAは各統計データの先頭LBAである。平均アクセス量は数式(2)で示す計算式によって計算される。アクセス総和およびコマンド重畳数は平均アクセス量を計算するために使用される。尚、表中の値は、例として図３のコマンド履歴の状態における値を示している。尚、統計データ領域402もコマンド履歴領域4

01と同様有限であるため、統計データ領域402に空きがなくなった場合、統計データでもっとも評価値の低い統計データを削除して領域を確保する。

【００１７】図５は、レジスタ群104の構成図の一例である。コマンド履歴に関する各レジスタ（504～508）と統計データに関する各レジスタ（509～513）は、レジスタの内容をアドレスデータとして出力することがあるのでアドレスマルチプレクサ503を介してアドレスバス502と接続されている。

【００１８】各レジスタの役割について説明する。コマンド履歴先頭レジスタ504はワークメモリ102上のコマンド履歴領域401の先頭を示す。コマンド履歴境界レジスタ505はコマンド履歴領域401の終端を示す。コマンド履歴最大数レジスタ506は、現在保持されているコマンド履歴の総数を示す。コマンド履歴カレントポインタ507は、コマンド履歴領域401の中で、現在のコマンド履歴情報の中でどの部分が先頭になるかを示す。これは、コマンド履歴領域がリングバッファになっているためである。コマンド履歴カウンタ508は、コマンド履歴情報を用いて処理を行う際に利用する。

【００１９】統計データ先頭レジスタ509はワークメモリ102上の統計データ領域402の先頭を示す。統計データ境界レジスタ510は統計データ領域402の終端を示す。統計データ最大数レジスタ511は、現在保持されている統計データの総数を示す。統計データカレントポインタ512は、統計データ領域512の中で、現在の統計データ領域の中でどの部分が先頭になるかを示す。統計データカウンタ513は、統計データ情報を用いて処理を行う際に利用する。

【００２０】コマンドレジスタ514は、現在のコマンドに関する情報を保持する。コマンド履歴レジスタ515は、ワークメモリ102より読み出したコマンド履歴情報を保持するために使用される。統計データレジスタ516は、ワークメモリ102より読み出した統計データを保持するために使用される。統計データLBAバッファレジスタ517は、統計データレジスタ516のLBA情報のみを保持するために使用される。平均アクセス数保持レジスタ518は、平均アクセス数を保持するために使用される。評価値保持レジスタ519は、評価値を保持するために使用される。統計LBA最小値レジスタ520は、統計データの中で最小のLBAの値を保持する。統計LBA最大値レジスタ521は統計データの中で最大のLBAの値を保持する。

【００２１】評価値変換テーブル522は、図６の評価値変換テーブル601に示すように、コマンド履歴番号と評価値を対応付けるための変換データが保持される。コマンド履歴番号と評価値は対になって記録されており、コマンド履歴番号をこのテーブルに入力すると、対応する評価値が出力される。このテーブルは必要に応じて動的に更新できるようにRAMで構成してもよい。また、このテーブルの更新をホストなど外部機器から行えるようにするために、テーブル制御用のコマンドセットを用意し、それを用いて外部から動的にテーブルの制御を行ってもよい。

【００２２】先読み量変換テーブル523は図６の先読み量変換テーブル602に示すように、統計データのアクセスLBAの範囲と先読み比率を対応付けるための変換データが保持される。統計データのアクセスLBAの範囲と先読み比率は対になって記録されており、統計データのアクセスLBAの範囲を入力すると先読み比率が出力される。このテーブルは必要に応じて更新できるようにRAMで構成してもよい。

【００２３】外部回路によって先読み予測装置101および201を制御する手順を説明する。図７は先読み予測装置の制御の手順を表したフローチャートである。これについて説明する。

【００２４】外部回路はまず、ステップ700を実行する。

【００２５】ステップ700では、ワークメモリ102のコマンド履歴カレントポインタ507の示す位置にホストからのコマンドデータを記録する。

【００２６】ステップ701では、コマンド履歴カレントポインタ507に1を加える。

【００２７】ステップ702では、コマンド履歴カレントポインタ507とコマンド履歴境界レジスタ505の値を比較し、コマンド履歴カレントポインタ507の値が大きい場合は、ステップ703を実行する。

【００２８】ステップ703では、コマンド履歴カレントポインタ507をクリアする。

【００２９】ステップ704では、加重統計回路103を起動して重み付け統計処理を行なう。動作の詳細については後述する。

【００３０】ステップ705では、先読み量計算回路105を起動して先読み量を算出する。動作の詳細については後述する。

【００３１】加重統計回路103の動作について説明する。図８及び図９は加重統計回路103が重み付け統計処理を行う際の動作を示すフローチャートである。

【００３２】加重統計回路103が外部回路によって起動されると、まず、ステップ801が実行され、コマンド履歴カウンタ508と統計データLBAバッファレジスタ517がクリアされる。

【００３３】ステップ802では、ワークメモリ102上のコマンド履歴カレントポインタ507の示す場所のコマンド履歴の値から統計データ形式405に基づいて初期統計データを作成する。具体的には、コマンド履歴番号を評価値変換テーブル601を用いて変換したものを評価値として、コマンド履歴のLBA をLBAとして、コマンド履歴のアクセスセクタ数を平均アクセス量とアクセス総和として、そしてコマンド重量数を1と設定する。このデータをワークメモリ102上の統計データ領域402の先頭に記録

する。

【００３４】ステップ803では、統計データ最大数レジスタ５１１に１をセットする。

【００３５】ステップ804では、重畳フラグをクリアする。この重畳フラグは、統計処理を行う際、現在処理対象になっているコマンドが統計データ領域の記録されているデータと重なる部分が存在したかを示す。

【００３６】ステップ805では、コマンド履歴カウンタ５０８へ１を足す。

【００３７】ステップ806では、ワークメモリ102のコマンド履歴領域401から、コマンド履歴カレントポインタ５０７にコマンド履歴カウンタ508を足した値の示す場所からコマンド履歴を取り出し、コマンド履歴レジスタ515に読み出す。

【００３８】ステップ807では、統計データカウンタ513をクリアする。

【００３９】ステップ808では、ワークメモリ102上の統計データ領域の空き領域を調べ、空き領域が存在する場合は、ステップ809を実行する。そうでない場合はステップ810へ制御を移す。

【００４０】ステップ809では、統計データカウンタ513と統計データ最大数レジスタ511の値を比較して統計データカウンタ513の値が最大数レジスタ511の値よりも大きい場合に、ステップ901へ制御を移す。そうでない場合は、ステップ810を実行する。

【００４１】ステップ810では、ワークメモリ102上の統計データ領域402内の統計データカレントポインタ512と統計データカウンタ513を足した値が示す領域の統計データを読み出し、統計データレジスタ516にセットする。

【００４２】ステップ811は、コマンド履歴レジスタ515と統計データレジスタ516の示す領域が重なっているかを比較し、重なっている場合はステップ812を実行し、そうでない場合はステップ820を実行する。

【００４３】ステップ812では、ステップ811の比較結果より、コマンド履歴と統計データの重なりが存在することが分かったため、重畳フラグをセットする。

【００４４】ステップ813では、統計データレジスタ516のコマンド重畳数に１を加える。

【００４５】ステップ814では、統計データレジスタ516のアクセス総和にコマンド履歴レジスタ515のセクタ数の値を加える。

【００４６】ステップ815では、統計データレジスタ516内のコマンド重畳数とアクセス総和から平均アクセス量を計算する。また、現在のコマンド履歴の番号を評価値変換テーブル601を用いて評価値に変換し、その値を統計データレジスタ516の評価値へ加える。

【００４７】ステップ816では、統計データのLBAとコマンド履歴のLBAを比較し、コマンド履歴のLBAが大きい場合は、ステップ906を実行する。そうでない場合は、ステップ817を実行する。

【００４８】ステップ817では、統計データレジスタ516の内容をワークメモリ102の統計データ領域402内の統計データカレントポインタ512と統計データカウンタ513の値を加えた値が示す領域へ記録する。

【００４９】ステップ818では、統計データカウンタ513へ1を足す。

【００５０】ステップ819では統計データLBAバッファレジスタ517に統計データレジスタ516の統計データLBAを設定し、ステップ808を実行する。

【００５１】ステップ820では、重畳フラグがセットされているかをチェックし、セットされている場合は、ステップ818を実行する。そうでない場合は、ステップ821を実行する。

【００５２】ステップ821では、統計データLBAバッファレジスタ517の値とコマンド履歴レジスタ515のLBAを比較し、コマンド履歴レジスタ515の値が大きい場合は、ステップ818を実行する。そうでない場合は、ステップ823を実行する。

【００５３】ステップ822では、統計データレジスタ516の値とコマンド履歴レジスタ505のLBAを比較し、コマンド履歴レジスタ505の値が大きい場合は、ステップ818を実行する。そうでない場合は、ステップ822を実行する。

【００５４】ステップ823では、統計データカレントポインタ512の示すデータ以降のデータをすべて後ろにシフトして、新しい統計データを挿入できる領域を用意する。そして、コマンド履歴レジスタ515の値から、ステップ802で作成した同様の形式で新規統計データを作成し、そのデータを挿入する。

【００５５】ステップ824では、統計データ最大数レジスタ511に1を足す。

【００５６】ステップ825では、重畳フラグをセットして、ステップ804を実行する。

【００５７】ステップ901では、重畳フラグをチェックする。セットされていればステップ904を実行する。セットされていなければステップ902を実行する。

【００５８】ステップ902はでは、コマンド履歴レジスタ515の値からステップ802で述べたのと同様の形式で、コマンド履歴レジスタ515から新規統計データを作成し、ワークメモリ102の統計データ領域402内で、現在存在する統計データの最後尾に追加する。

【００５９】ステップ903では、統計データ最大数レジスタ511に1を加える。

【００６０】ステップ904では、コマンド履歴カウンタ５０８に１を加える。

【００６１】ステップ905では、コマンド履歴カウンタ５０８とコマンド履歴最大数レジスタ506の内容を比較し、コマンド履歴カウンタ508の値がLBAが大きければ、処理を終了する。そうでなければ、ステップ804へ制御を移す。

【００６２】ステップ906では、統計データレジスタ516のLBAの値をコマンド属性レジスタ515のLBAの値に置き換える。

【００６３】ステップ907では、統計データ最大数レジスタ511に１を加える。

【００６４】ステップ908では、統計データカレントポインタ512の示すデータの次のデータ以降のすべてのデータをすべて後ろにシフトして、統計データレジスタ516の内容を新規統計データとして挿入する。

【００６５】ステップ909では、統計データカウンタ513に2を加え、ステップ819を実行する。

【００６６】ステップ910では、統計データカウンタ513の値を一時的に退避し、統計データカウンタ513をクリアする。

【００６７】ステップ911では、統計データカレントポインタ512に統計データ先頭レジスタの値をセットする。

【００６８】ステップ912では、ワークメモリ102１から、統計データを統計データレジスタ516に読み出し、統計データの評価値を評価値保持レジスタ519へ保存する。

【００６９】ステップ913では、統計データカウンタ513へ１を加える。

【００７０】ステップ914では、統計データレジスタ516へ、統計データを読みむ。

【００７１】ステップ915では、評価値保持レジスタ519と統計データレジスタ516の値を比較し、統計データレジスタ516の値が大きければ、ステップ916を実行する。

【００７２】ステップ916では、評価値保持レジスタ519に統計データレジスタ516の値の評価値を記録し、統計データカレントポインタ512の値を読み出した統計データの位置へセットする。

【００７３】ステップ917では、統計データカウンタ513の値と統計データ最大数を比較し、等しければ、ステップ918を実行する。そうでなければ、ステップ908へ制御を移す。

【００７４】ステップ918では、ワークメモリ102の統計データカレントポインタ507の示す位置のデータを削除する。

【００７５】ステップ919では、統計データ最大数レジスタ511から1を引く。

【００７６】ステップ920では、統計データカウンタを復旧し、ステップ809へ制御を移す。

【００７７】加重統計回路103が統計データの生成を終了すると、次に、統計データのLBAの範囲を求めるために、加重統計回路103は図１０で示すフローチャートに従った動作を行う。これについて説明する。

【００７８】まず、ステップ1001が実行され統計データカウンタ513がクリアされる。

【００７９】ステップ1002では、統計データをワークメモリ102の統計データ領域402から統計データレジスタ516に読み出す。

【００８０】ステップ1003では統計LBA最大値レジスタ521と統計LBA最小値レジスタ520に統計データレジスタ516の値をセットする。

【００８１】ステップ1004では、統計データカウンタ513に1を加える。

【００８２】ステップ1005では、統計データレジスタ516に統計データをワークメモリ102から読み出す。

【００８３】ステップ1006では、統計データレジスタ516のLBAと統計LBA最大値レジスタ521の値を比較し、統計データのLBAが大きい場合は、ステップ1008を実行する。そうでない場合は、ステップ1007を実行する。

【００８４】ステップ1007では、統計データレジスタ516のLBAと統計LBA最小値レジスタ520の値を比較し、統計データのLBAが小さい場合は、ステップ1009を実行する。そうでない場合は、ステップ1010を実行する。

【００８５】ステップ1008では、統計データレジスタ516のLBAの値を統計LBA最大値レジスタ521へセットする。

【００８６】ステップ1009では、統計データレジスタ516のLBAの値を統計LBA最小値レジスタ520へセットする。

【００８７】ステップ1010では、統計データカウンタ513と統計データ最大数レジスタ521の値を比較し、統計データカウンタ513の値が大きければ、処理を終了する。そうでない場合は、ステップ1004を実行する。

【００８８】図１１は先読み計算回路105の動作のフローチャートを示したものである。これについて説明する。

【００８９】先読み量計算回路が外部回路によって起動されると、まずステップ1100が実行される。このステップでは、ホストからデータバス111を介して転送されてきたLBAとセクタ数をコマンドレジスタ514に記録する。

【００９０】ステップ1101では、統計データカウンタ513と評価値保持レジスタ519および平均アクセス数保持レジスタ518をクリアする。

【００９１】ステップ1102では、統計データカウンタ513と統計データ最大数レジスタ511の値を比較し、もし統計データカウンタ513の値が大きければ、ステップ1109を実行する。そうでなければ、ステップ1103を実行する。

【００９２】ステップ1103では、ワークメモリ102から統計データを統計データレジスタ516へ読み出す。

【００９３】ステップ1104では、統計データレジスタ516とコマンドレジスタ514の値を比較し、アクセス領域の重なりを調べる。アクセス領域の重なりがあれば、ステップ1105を実行する。そうでなければステップ1108を実行する。

【００９４】ステップ1105では、統計データレジスタ516の評価値と評価値保持レジスタ519の値を比較し、統計データレジスタ516の値が大きければ、ステップ1106を

実行する。そうでなければ、ステップ1108を実行する。

【００９５】ステップ1106では、評価値保持レジスタ519の値を統計データレジスタ516の評価値に置きかえる。

【００９６】ステップ1107では、平均アクセス数保持レジスタ518の値を、統計データレジスタ516の平均アクセス数に置きかえる。

【００９７】ステップ1108では、統計データカウンタに1を加える。

【００９８】ステップ1109では、統計LBA最大値レジスタ521から統計LBA最小値レジスタ520の値を引くことで統計データのアクセス範囲を求める。

【００９９】ステップ1110では、アクセス幅の値を先読み量変換テーブル602を用いて先読み比率データに変換する。

【０１００】ステップ1111では、先読み比率と平均アクセス数の積を取ることで先読み量を算出する。

【０１０１】次に、本発明を用いて構成されたハードディスクについて説明する。

【０１０２】(1)ディスク制御装置1203内に先読み予測装置1211を内蔵する場合

図１２は、本発明を用いて構成されたハードディスクを用いた情報器機の構成図の一例である。ホスト1200はホストインタフェース1201によってハードディスク1202と接続されている。ハードディスク1202はディスク制御装置1203、キャッシュメモリ1220、ROM1208、CPU1213、サーボ系制御回路1224、そして記録媒体1228から構成されている。ディスク制御装置1203は、インタフェース制御回路1204、シーケンサ1207、先読み予測装置1211、アドレスマルチプレクサ1209および1214、データマルチプレクサ1222、ECC1227、ディスクフォーマッタ1226から構成されている。キャッシュメモリ1220は、ハードディスク1202とホスト1200間で送受を行うデータを記憶しておくと共に、キャッシュメモリ1220内のデータを管理するためのセグメント管理テーブルなどの情報を記憶する場合もある。

【０１０３】また、先読み予測装置1211が図１で示した先読み予測装置101のようにワークメモリ102を内蔵せず、図2で示した先読み予測装置201の構成になっている場合には、先読み予測装置1211のワークメモリの働きもする。ROM1206はハードディスク1202の制御に必要なファームウェア、制御パラメータなどが記憶されている。CPU1213は、ファームウェアの実行、サーボ系制御回路1224の制御に用いられる。サーボ系制御回路1224は記録媒体1228の機械系とそれに関連のある部分の制御を行う。記録媒体1228はデータを記録する磁気ディスク、磁気ディスクからの信号と、ディスクフォーマッタとの間のデータ変換を行い読み出し書き込みを行うためのR/Wチャネル、磁気ディスクを制御する機械系等から構成されている。尚、記録媒体としては磁気ディスク以外の記録媒体でも良い。

【０１０４】インタフェース制御回路1204は、ホスト1200とインタフェース1201を介して接続されており、ホスト1200とハードディスク1202間の実際のデータ転送を制御し、必要なデータを内部データバス1212へ載せる。シーケンサ1207はインタフェース制御信号1205、モジュール制御信号1210を用いて、ディスク制御装置1203内の各モジュールおよびキャッシュメモリ1220を制御する。また、必要に応じてCPU割り込み制御信号1206によってCPU1213へ割り込みを発生させる。アドレスマルチプレクサ1209はシーケンサ1207の内部アドレスバス1215とCPU1213のCPUアドレスバス1219を切り替えて、先読み予測装置1211のアドレスバスと接続する。先読み予測装置1211はホストからコマンドが来た際にそのコマンドをコマンド履歴として保持し、コマンド履歴に関する統計処理を行い、現在のコマンドにおける評価値と先読み量を算出する。

【０１０５】また、ディスクフォーマッタ制御信号1221が存在する場合は、この信号によって直接ディスクフォーマッタに先読み量を設定し、先読みを制御する。アドレスマルチプレクサ1214は、キャッシュメモリ1220と接続するディスク制御装置1203内の内部アドレスバス1215とCPUアドレスバス1219及びディスクフォーマッタアドレスバス1216及び、先読み予測装置1211がキャッシュメモリ1220をワークメモリとして用いる場合に使用されるアドレスバス1231を切り替える。データマルチプレクサ1222はCPUデータバス1218及びフォーマッタデータバス1223とディスク制御装置1203内の内部データバス1212との接続を制御する。ECC1227はディスクフォーマッタ1226とECCデータバス1229を介して接続されており、記録媒体1228より読み出したデータのエラー訂正を行う。ディスクフォーマッタ1226は、信号バス1230を介して記録媒体1228との間でデータ転送を行う。

【０１０６】本実施例では、キャッシュメモリ1220の制御方式として可変セグメント方式を用いた。ただし、先読み予測装置1211自体は、キャッシュメモリの制御方式に依存せずに使用することが可能である。

【０１０７】可変セグメント方式について説明する。図１４にこの方式におけるキャッシュメモリ1400内の状態とセグメント管理テーブル1401を示す。キャッシュメモリ1400はワークエリアとデータエリアに分けて使用される。ワークエリアには、ファームウェアの情報およびセグメント管理テーブル、場合によっては先読み予測装置1211のデータが記憶される。ただし、セグメント管理テーブルは高速なアクセスが要求されるので、キャッシュメモリ1400以外の場所（ディスク制御装置1203内、CPU1213内の作業用不揮発性メモリ等）の記録領域に記録される場合もある。データエリアはハードディスク1202とホスト1200間で送受されるデータが記憶される。

【０１０８】本方式では、データエリア内のLBAの連続するデータを一まとまりとして セグメントとし、各セグ

メントをセグメント管理テーブルで管理する。セグメント管理テーブルの内容は、図１４のセグメント管理テーブル1401に示すように、そのセグメントの評価値、各セグメント内のデータの先頭LBA、キャッシュメモリ1220内でのデータの先頭アドレス、セグメントのデータサイズが記録されている。新しくデータを記憶する場合、データアクセスポインタ1402の示すアドレスからデータを記憶させ、データの記憶が終了した時点で、データの記憶を開始した先頭アドレスおよび対応するLBA、そして記憶したデータのサイズをセグメント管理テーブルに登録し、生成したセグメントの最後尾の次のアドレス値にデータアクセスポインタ1402の値を更新する。データエリアはリングバッファになっており、データエリアの最後までデータが記憶されると、データエリアの先頭に戻って後続のデータを記憶させる。この時、新しいデータによって上書きされるセグメントが存在する場合、そのセグメントを破棄し、セグメント管理テーブルの対応するセグメント情報を無効にする。

【０１０９】ホストからコマンド受けた場合についてハードディスク1202の動作について説明する。

【０１１０】図１５は、ディスク制御装置1203とCPU1213の動作を示すフローチャートである。ステップ1502に示すようにディスク制御装置1203がホスト1200からコマンドを受けると、ディスク制御装置1203はコマンド受信割り込みをCPU1213に対して発行しCPU1213を呼び出す。次にディスク制御装置1203は、ステップ1503に示すように、先読み予測装置1211へホスト1200からのコマンドデータを転送する。その後、ステップ1504では、先読み予測装置1211が統計処理と先読み量の算出を行う。そして、ステップ1505で、先読み予測装置1211は、直接ディスクフォーマッタ1226に対して最大先読み量を設定する。

【０１１１】一方、CPU1213は、ディスク制御装置1203からコマンド受信割り込みを受けると、ステップ1512に示すようにコマンドの解析を行う、そして、ステップ1513では、ホスト1200からのコマンドがリードコマンドかどうか判定し、違えば、ステップ1516で示すように、他のコマンドの処理をディスク制御装置1203を操作して行う。リードコマンドなら、ステップ1515で、キャッシュメモリの内のデータとのヒット判定を行う。データがヒットした場合、ステップ1506に示すように、ディスク制御装置1203を制御してキャッシュメモリ1220上のヒットしたデータをホスト1200へ転送する。またその時点で行われている先読みに関する処理はそのまま継続する。

【０１１２】ヒットしなかった場合、ステップ1516でハーフヒット判定を行う。

【０１１３】ミスの場合、ステップ1517を実行する。このステップでは、新しいセグメントをキャッシュメモリ1220上に生成する。もしキャッシュメモリ1220に空き領域がない場合、セグメント管理テーブル1401の各セグメントの評価値とデータサイズを比較して、データサイズが平均アクセス数保持レジスタ518の値よりも大きく、その中で評価値が最も低いセグメントを調べ、そのセグメントの先頭にデータアクセスポインタ1402を設定する。そのセグメント情報を削除する。新たにデータアクセスポインタ1402の値を先頭アドレスとし、評価値保持レジスタ519の値を評価値として設定した新規セグメントを生成する。

【０１１４】ハーフヒットの場合、ステップ1518を実行する。このステップでは、評価値保持レジスタ519の値が高ければ、ステップ1519でセグメント管理テーブル1401内のハーフヒットしたセグメントのセグメントサイズを更新する。この更新方法には以下に示す2通りがある。

【０１１５】(a) ハーフヒットの場合でハーフヒットしたセグメントのデータが図１７の1700に示すようにホスト要求データの後ろにあたる場合、データアクセスポインタ1402を図１７の1700から1701で示す位置に更新する。これによってホスト要求データを記録媒体1228から読み出した場合、後ろの部分がちょうどハーフヒットしたセグメントのデータと重なるようにデータアクセスポインタ1402を設定される。そしてセグメント管理テーブル1401内のハーフヒットしたセグメントのセグメントサイズを更新する。

【０１１６】(b) ハーフヒットの場合でハーフヒットしたセグメントのデータが図１８の1800に示すように、ホスト要求データの前半部にあたる場合、データアクセスポインタ1402を図１８の1800から1801で示す位置に更新する。これによって不足分のデータのみをハーフヒットしたセグメントの直後から連続して読み出すようにする。そしてセグメント管理テーブル1401内のハーフヒットしたセグメントのセグメントサイズを更新する。

【０１１７】ミスおよびハーフヒットの場合ステップ1507に示すようにCPU1213がディスク制御装置1203を操作して先読みを起動する。

【０１１８】ステップ1508で、ホスト要求分のデータがキャッシュメモリ1220に読みこまれたかを判定し、読みこまれた場合、ディスク制御装置1203はCPU1213に対してセグメント更新割り込みを発行する。CPU1213は、ステップ1519でセグメント情報を更新し、ディスク制御装置1203を起動して、ステップ1506を実行させる。次にディスク制御装置1203は、ステップ1509で先読みを継続し、ステップ1510で最大先読み量に先読み量が達したか判定する。最大先読み量に先読み量が達した場合、ステップ1511で先読みを停止させ、CPU1213に先読み停止割り込みを発行する。これを受けたCPU1213はステップ1520でセグメント情報を更新する。

【０１１９】(2)ディスク制御装置1301の外部に先読み予測装置1302を搭載する場合

図１３は、本発明を用いて構成されたハードディスクを

用いた情報器機の構成図の一例である。図１２の構成との変更点は、先読み予測装置1302がディスク制御装置1301の外部に設置されている。このため、先読み予測装置1302はCPUデータバス1218とCPUアドレスバス1219およびCPU制御信号1217に接続されている。

【０１２０】ホストからコマンドを受けた場合についてハードディスク1300の動作について説明する。

【０１２１】図１６は、ディスク制御装置1301とCPU1213の動作を示すフローチャートである。ステップ1600に示すようにホストからコマンドを受けると、ディスク制御装置は1301はコマンド受信割り込みを発行してCPU1213を呼び出す。CPU1213は、ディスク制御装置1203からの割り込みを受けると、ステップ1608でコマンドの解析を行う、そして、ステップ1609では、ホスト1200からのコマンドがリードコマンドかどうか判定し、違えばステップ1601で、他のコマンドの処理をディスク制御装置1203を操作して行う。リードコマンドなら、ステップ1610で先読み予測装置1302へホスト1200からのコマンドデータを転送する。そして、ステップ1611で先読み予測装置1302を制御して統計処理と先読み量を算出する。この後、ステップ1612で、キャッシュメモリ1220内のデータのヒット判定を行う。データがヒットした場合、ステップ1602で、ディスク制御装置1301を制御してキャッシュメモリ1220上のヒットしたデータをホスト1200へ転送する。先読み関しては特に制御をせず現状を維持する。

【０１２２】ヒットしなかった場合、ステップ1613でハーフヒット判定を行う。

【０１２３】ミスの場合、ステップ1614で、新しいセグメントをキャッシュメモリ1220上に生成する。もしキャッシュメモリ1220に空き領域がない場合、セグメント管理テーブル1401の各セグメントの評価値とデータサイズを比較して、データサイズが平均アクセス数保持レジスタ518の値よりも大きく、その中で評価値が最も低いセグメントを調べ、そのセグメントの先頭にデータアクセスポインタ1402を設定する。そして、セグメント情報を削除し、新たにデータアクセスポインタ1402の値を先頭アドレスとし、評価値保持レジスタ519の値を評価値として設定した新規セグメントを生成する。

【０１２４】ハーフヒットの場合、ステップ1615を実行する。このステップでは、評価値保持レジスタ519の値が高ければ、ステップ1615でセグメント管理テーブル1401内のハーフヒットしたセグメントのセグメントサイズを更新する。この更新方法には以下に示す2通りがある。

【０１２５】(a) ハーフヒットの場合でハーフヒットしたセグメントのデータが図１７の1700に示すようにホスト要求データの後ろにあたる場合、データアクセスポインタ1402を図１７の1700から1701で示す位置に更新する。これによって次ホスト要求データを記録媒体1228から読み出した場合、後ろの部分がちょうどハーフヒットしたセグメントのデータと重なるようにデータアクセスポインタ1402を設定される。そしてセグメント管理テーブル1401内のハーフヒットしたセグメントのセグメントサイズを更新する。

【０１２６】(b) ハーフヒットの場合でハーフヒットしたセグメントのデータが図１８の1800に示すように、ホスト要求データの前半部にあたる場合、データアクセスポインタ1402を図１８の1800から1801で示す位置に更新する。これによって不足分のデータのみをハーフヒットしたセグメントの直後から連続して読み出すようにする。そしてセグメント管理テーブル1401内のハーフヒットしたセグメントのセグメントサイズを更新する。

【０１２７】ミスおよびハーフヒットの場合、ステップ1601を実行する。

【０１２８】このステップでは、先読み予測装置1302から最大先読み量を読み出し、ディスク制御装置1301に指示をだし、ディスクフォーマッタ1304に最大先読み量を設定する。そして、ステップ1603に示すように先読みを起動する。

【０１２９】ステップ1604で、ホスト1200要求分のデータがキャッシュメモリ1220に読みこまれたかを判定し、読みこまれた場合、ディスク制御装置1301はCPU1213に対してセグメント更新割り込みを発生させる。CPU1213は、ステップ1617でセグメント情報を更新し、ディスク制御装置1301を起動して、ステップ1602を実行させる。次にディスク制御装置1301は、ステップ1605で先読みを継続し、ステップ1606で最大先読み量に先読み量が達したか判定する。最大先読み量に先読み量が達した場合、ステップ1607で先読みを停止させ、CPU1213に対してディスク制御装置1301は先読み停止割り込みを発行する。これを受けたCPU1213はステップ1618でセグメント情報を更新する。

【０１３０】本発明を用いて構成されたディスクアレイコントローラの一例を図１９に示す。インタフェース制御回路1901は、ホスト1200とのインタフェース1201を制御する。マイクロプロセッサ1903は、ディスクアレイコントローラ1900内の各モジュールを制御する。記録装置制御回路1905は、記録装置1906を制御する。キャッシュメモリ1902は、ホスト1200及び記録装置1906から転送されて来たデータを一時的に保持する。バス1901は、ディスクアレイコントローラ1900内の各モジュールを接続するために使用する。記録装置接続バス1908は、ディスクアレイコントローラ1900と記録装置1906を接続するために使用する。

【０１３１】ホスト1200からコマンドを受けた場合についてディスクアレイ1900の動作について説明する。

【０１３２】図２０は、ディスクアレイコントローラ1900の全体の動作とその内部のマイクロプロセッサ1903の動作を示すフローチャートである。このフローチャートで示す動作は図１６で示すフローチャートと比較して、

ステップ2016で最大先読み量を設定するのが記録装置制御回路1905である以外同様である。本実施例では、磁気ディスクを記録媒体に用いた場合について説明したが、無論、記録媒体が光磁気ディスクなどの他の円盤状記録媒体を用いた場合についても同様である。

【０１３３】

【発明の効果】本発明では、ホストからのコマンドに対して、その履歴を保持し、記録媒体上の各領域の平均アクセス量を求める。以後その領域に対してホストからアクセスがあった場合、先読み量を最適化し、キャッシュメモリへ余分なデータを読み込まないように先読みを制御する。このためキャッシュメモリの使用効率が改善すると共に、必要なデータを読み込んだ時点でディスクフォーマッタの動作が停止しているため、ホストの要求から、シーク動作の開始までの時間を短縮することが可能である。

【０１３４】また、コマンド履歴の時間的な情報を考慮した重み付け統計処理を行うことで、記録媒体上の各領域のホストからのアクセス頻度を求め、キャッシュメモリの各セグメントが対応する記録媒体上の領域についての評価値を求めておく、そして、ホストからアクセスがあった場合、その評価値に基づいてキャッシュメモリ内への記録媒体からのデータの読みこみ位置、およびデータの置換を制御することで、キャッシュメモリの使用効率を改善できる。

【図面の簡単な説明】

【図１】本発明を用いて構成された先読み予測装置の構成図である。

【図２】図１の先読み予測装置においてワークメモリを外部に設けた場合の構成図である。

【図３】先読み予測装置で行われる重み付け統計処理の概略図である。

【図４】ワークメモリの構成と各領域でのデータ形式および統計処理を行う際に使用される計算式の説明図である。

【図５】図１のレジスタ群102の詳細な構成図である。

【図６】図５の各テーブルの変換形式の説明図である。

【図７】先読み予測装置の動作を説明したフローチャートである。

【図８】加重統計回路が重み付け統計処理を行う際の動作を説明したフローチャートである。

【図９】加重統計回路の重み付け統計処理を行う際の動作を説明したフローチャートである。

【図１０】加重統計回路が統計データのLBAの範囲を求める際の動作を説明したフローチャートである。

【図１１】先読み量計算回路の動作を説明したフローチャートである。

【図１２】図１あるいは図２の先読み予測装置を内蔵したディスク制御装置を用いて構成されたハードディスク装置の構成図である。

【図１３】図１あるいは図２の先読み予測装置をディスク制御装置の外部に設けて構成されたハードディスク装置の構成図である。

【図１４】図１２あるいは図１３のハードディスクにおけるキャッシュメモリの制御方式の説明図である。

【図１５】図１２のハードディスクがホストからコマンドを受けた場合の動作を説明したフローチャートである。

【図１６】図１３のハードディスクがホストからコマンドを受けた場合の動作を説明したフローチャートである。

【図１７】図１５および図１６のハードディスクにおいてキャッシュメモリのデータに対してハーフヒットが起こった場合の処理の説明図である。

【図１８】図１５および図１６のハードディスクにおいてキャッシュメモリのデータに対してハーフヒットが起こった場合の処理の説明図である。

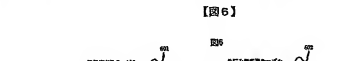
【図１９】図１あるいは図２の先読み予測装置を搭載したディスクアレイコントローラを用いて構成されたディスクアレイの構成図である。

【図２０】図１９のディスクアレイがホストからコマンドを受けた場合の動作を説明したフローチャートである。

【符号の説明】

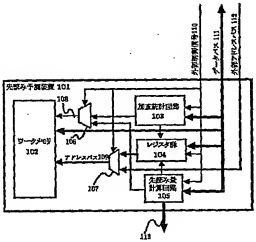
106…制御信号マルチプレクサ、107…アドレスマルチプレクサ、108…ワークメモリ制御信号、404…コマンド領域データ形式、405…統計データ領域データ形式、503…アドレスマルチプレクサ、1205…インタフェース制御信号、1210…モジュール制御信号、1216…ディスクフォーマッタアドレスバス、1223…ディスクフォーマッタデータバス、1231…アドレスバス、1221…ディスクフォーマッタ制御信号、1229…ECCデータバス、1230…信号バス、1225…記録媒体制御信号、1907…バス、1908…記録装置制御バス

【図６】



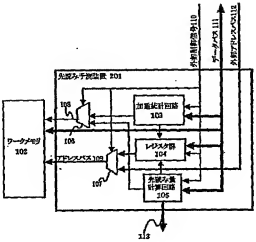
【図１】

図1



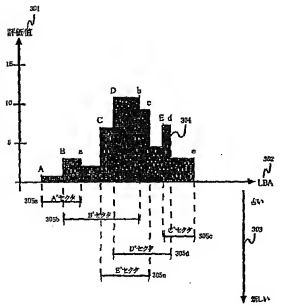
【図２】

図2



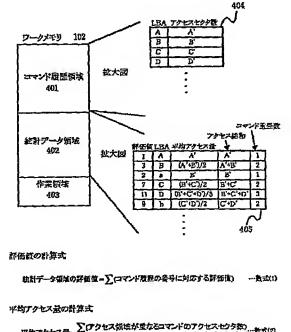
【図３】

図3



【図４】

図4



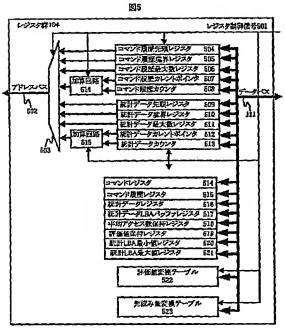
評価値の計算式

統計データ領域の評価値＝$\sum$(コマンド履歴の各号に対応する評価値)　　…数式(1)

平均アクセス量の計算式

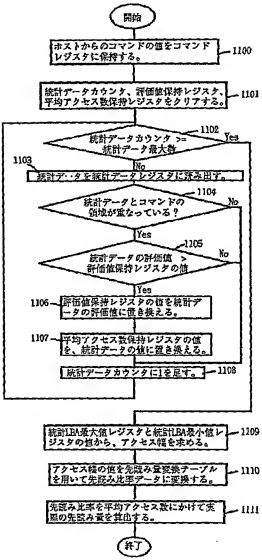
平均アクセス量＝$\frac{\sum(\text{アクセス領域が重なるコマンドのアクセスセクタ数})}{\text{アクセス領域が重なるコマンド数}}$　…数式(2)

【図５】


B05
レジスタ群104
レジスタ制御信号501
コマンド指示元識レジスタ　B04
コマンド履歴保持レジスタ　B05
コマンド投入数レジスタ　B06
コマンド起動コンパレータ　B07
コマンド起動カウンタ　B08
統計データ元取レジスタ　B09
統計データ差異レジスタ　B10
統計データ最大数レジスタ　B11
統計データカットポイント　B12
統計データカウンタ　B13
コマンドレジスタ　B14
コマンド識別レジスタ　B15
統計データレジスタ　B16
統計データLBAバックアップレジスタ　B17
平均アクセス数保持レジスタ　B18
評価値変換用レジスタ　B19
統計LBA最小値レジスタ　B20
統計LBA最大値レジスタ　B21
評価値変換テーブル　B22
先読み量変換テーブル　B23
アドレスバス　B02
データバス　B11
制御回路　B14
制御回路　B15
B03

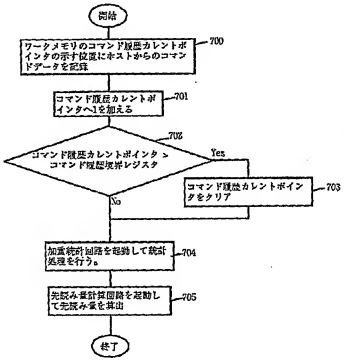
【図１１】

図11


開始
ホストからのコマンドの値をコマンドレジスタに保持する。
1100
統計データカウンタ、評価値保持レジスタ、平均アクセス数保持レジスタをクリアする。
1101
統計データカウンタ >= 統計データ最大数
1102
Yes
No
1103
統計データを統計データレジスタに読み出す。
統計データとコマンドの領域が変なっている？
1104
No
Yes
統計データの評価値 > 評価値保持レジスタの値
1105
No
Yes
1106
評価値保持レジスタの値を統計データの評価値に置き換える。
1107
平均アクセス数保持レジスタの値を、統計データの値に置き換える。
統計データカウンタに1を足す。
1108
統計LBA最大値レジスタと統計LBA最小値レジスタの値から、アクセス幅を求める。
1109
アクセス幅の値を先読み量変換テーブルを用いて先読み比率データに変換する。
1110
先読み比率を平均アクセス数にかけて実際の先読み量を算出する。
1111
終了

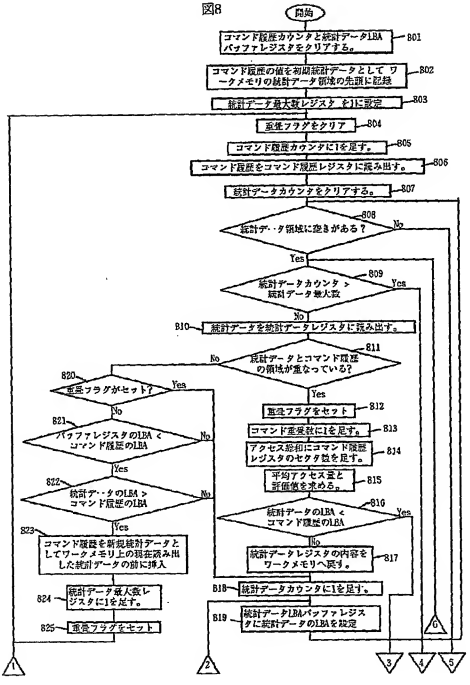
【図７】

図7


開始
ワークメモリのコマンド履歴カレントポインタの示す位置にホストからのコマンドデータを記録
700
コマンド履歴カレントポインタへ1を加える
701
コマンド履歴カレントポインタ＞コマンド履歴境界レジスタ
702
Yes
No
コマンド履歴カレントポインタをクリア
703
加重統計回路を起動して統計処理を行う。
704
先読み量計算回路を起動して先読み量を算出
705
終了

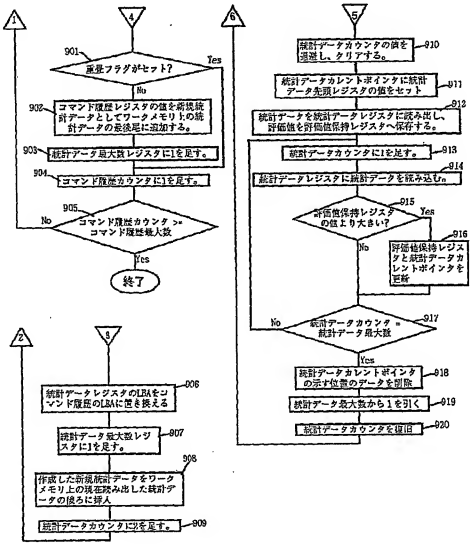
【図8】

図8


開始
コマンド履歴カウンタと統計データLBA バッファレジスタをクリアする。
801
コマンド履歴の値を初期統計データとして ワークメモリの統計データ領域の先頭に記録
802
統計データ最大数レジスタ a に設定
803
重要フラグをクリア
804
コマンド履歴カウンタに1を足す。
805
コマンド履歴をコマンド履歴レジスタに読み出す。
806
統計データカウンタをクリアする。
807
808
統計データ領域に空きがある？
No
Yes
809
統計データカウンタ > 統計データ最大数
Yes
No
810
統計データを統計データレジスタに読み出す。
811
統計データとコマンド履歴の領域が重なっている？
No
Yes
820
重要フラグがセット？
Yes
No
812
重要フラグをセット
813
コマンド重要数に1を足す。
814
アクセス総和にコマンド履歴レジスタのセクタ数を足す。
815
平均アクセス量と評価値を求める。
816
統計データのLBA < コマンド履歴のLBA
Yes
No
817
統計データレジスタの内容をワークメモリへ戻す。
818
統計データカウンタに1を足す。
819
統計データLBAバッファレジスタに統計データのLBAを設定
821
バッファレジスタのLBA < コマンド履歴のLBA
No
Yes
822
統計データのLBA > コマンド履歴のLBA
No
Yes
823
コマンド履歴を新規統計データとしてワークメモリ上の現在読み出した統計データの前に挿入
824
統計データ最大数レジスタに1を足す。
825
重要フラグをセット
1
2
3
4
5
6

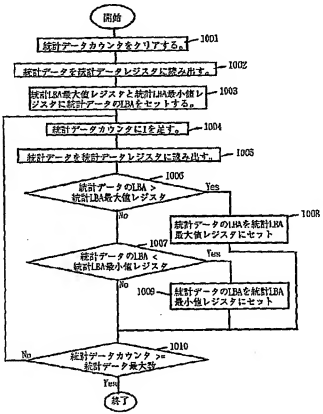
【図9】

図9


1
4
901
履歴フラグがセット?
Yes
No
902
コマンド履歴レジスタの値を前後統計データとしてワークメモリ上の統計データの最後尾に追加する。
903
統計データ最大数レジスタに1を足す。
904
コマンド履歴カウンタに1を足す。
905
No
コマンド履歴カウンタ >= コマンド履歴最大数
Yes
終了
6
5
統計データカウンタの値を退避し、クリアする。
910
統計データカレントポインタに統計データ先頭レジスタの値をセット
911
912
統計データを統計データレジスタに読み出し、評価値を評価値保持レジスタへ保存する。
統計データカウンタに1を足す。
913
914
統計データレジスタに統計データを読み込む。
915
評価値保持レジスタの値より大きい?
Yes
No
916
評価値保持レジスタと統計データカレントポインタを更新
No
917
統計データカウンタ = 統計データ最大数
Yes
統計データカレントポインタの示す位置のデータを削除
918
統計データ最大数から1を引く
919
統計データカウンタを復旧
920
2
3
統計データレジスタのLBAをコマンド履歴のLBAに置き換える
906
統計データ最大数レジスタに1を足す。
907
908
作成した新規統計データをワークメモリ上の現在読み出した統計データの後ろに挿入
統計データカウンタに2を足す。
909

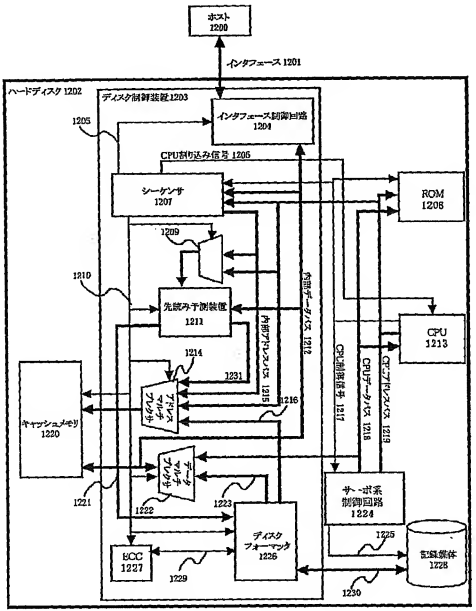
【図１０】

図10


開始
統計データカウンタをクリアする。
1001
統計データを統計データレジスタに読み出す。
1002
統計LBA最大値レジスタと統計LBA最小値レジスタに統計データのLBAをセットする。
1003
統計データカウンタに1を足す。
1004
統計データを統計データレジスタに読み出す。
1005
統計データのLBA > 統計LBA最大値レジスタ
1006
Yes
No
統計データのLBAを統計LBA最大値レジスタにセット
1008
統計データのLBA < 統計LBA最小値レジスタ
1007
Yes
No
統計データのLBAを統計LBA最小値レジスタにセット
1009
統計データカウンタ >= 統計データ最大数
1010
No
Yes
終了

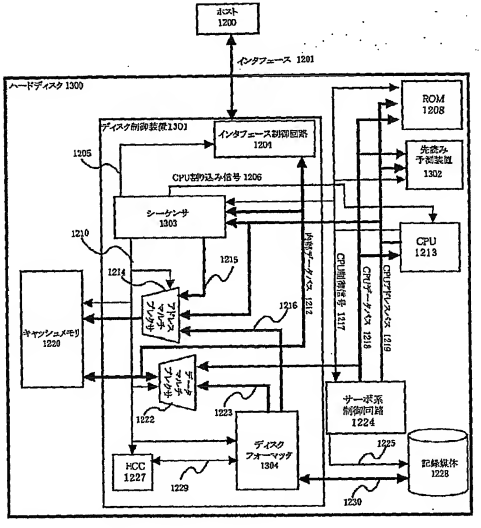
【図１２】

図12


ホスト
1200
インタフェース 1201
ハードディスク 1202
ディスク制御装置 1203
インタフェース制御回路
1204
1205
CPU割り込み信号 1206
シーケンサ
1207
ROM
1208
1209
1210
先読み予測装置
1211
内部データバス 1212
CPU
1213
1214
1231
内部アドレスバス 1215
CPUアドレスバス 1218
CPU制御信号 1217
CPUデータバス 1219
1216
キャッシュメモリ
1220
1221
1222
1223
サーボ系
制御回路
1224
ディスク
フォーマッタ
1226
1225
記録媒体
1228
ECC
1227
1229
1230

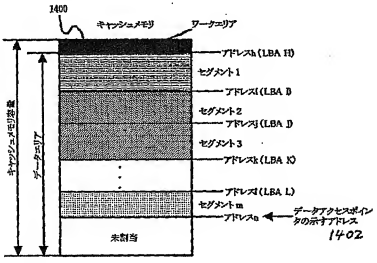
【図１３】

図13


ホスト
1200
インタフェース 1201
ハードディスク 1300
ディスク制御装置1301
インタフェース制御回路
1201
ROM
1208
先読み
予測装置
1302
1205
CPU制御込み信号 1206
シーケンサ
1303
1210
CPU
1213
1214
1215
内部データバス 1216
CPU制御信号 1217
CPUデータバス 1218
CPUアドレスバス 1219
キャッシュメモリ
1220
1216
1222
1223
サーボ系
制御回路
1224
ディスク
フォーマッタ
1304
1225
記録媒体
1228
HCC
1227
1229
1230

【図１４】

図14



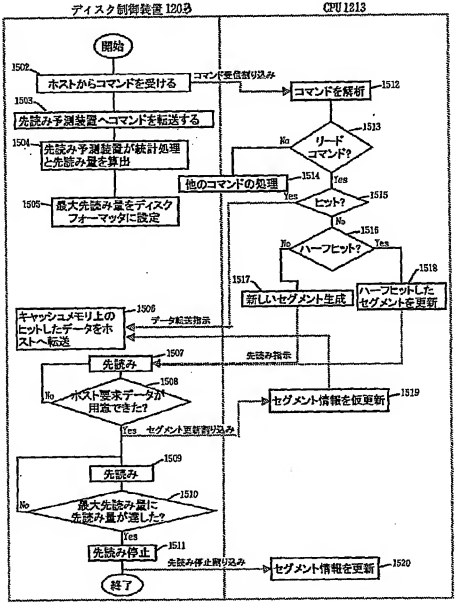
1401

セグメント管理テーブル

セグメント	評価値	対応するLBA	先頭アドレス	データサイズ
1	o	H	h	i-h
2	p	I	i	j-i
3	q	J	j	k-j
・	・	・	・	・
・	・	・	・	・
・	・	・	・	・
m	r	M	m	n-m
n	s	未使用	未使用	未使用

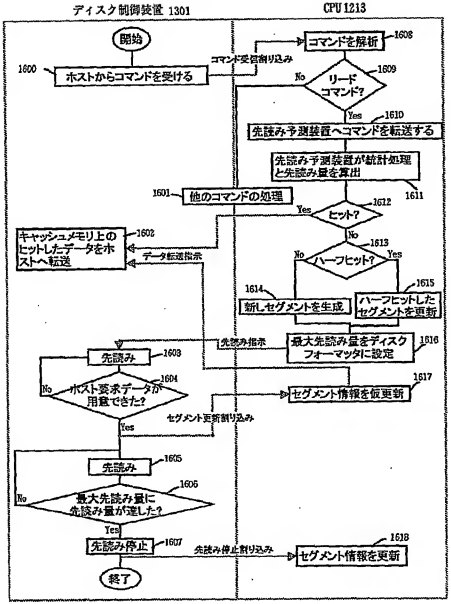
【図１５】

図15


ディスク制御装置 1203
CPU 1213
開始
1502
ホストからコマンドを受ける
コマンド受信割り込み
コマンドを解析
1512
1503
先読み予測装置へコマンドを転送する
1504
先読み予測装置が統計処理と先読み量を算出
No
リードコマンド?
1513
Yes
他のコマンドの処理
1514
1505
最大先読み量をディスクフォーマッタに設定
Yes
ヒット?
1515
No
No
ハーフヒット?
1516
Yes
1517
新しいセグメント生成
1518
ハーフヒットしたセグメントを更新
キャッシュメモリ上のヒットしたデータをホストへ転送
1506
データ転送指示
先読み
1507
先読み指示
No
ホスト要求データが用意できた?
1508
セグメント情報を仮更新
1519
Yes
セグメント更新割り込み
先読み
1509
No
最大先読み量に先読み量が達した?
1510
Yes
先読み停止
1511
先読み停止割り込み
セグメント情報を更新
1520
終了

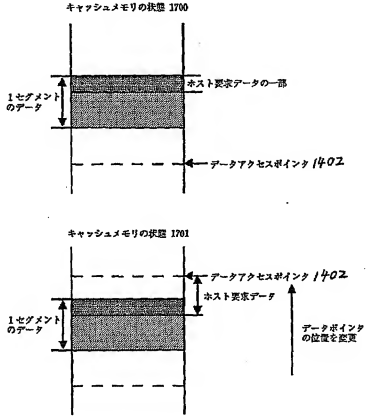
【図１６】

図16


ディスク制御装置 1301
CPU 1213
開始
コマンドを解析
1608
コマンド受信割り込み
1600
ホストからコマンドを受ける
No
リードコマンド?
1609
Yes
1610
先読み予測装置へコマンドを転送する
先読み予測装置が統計処理と先読み量を算出
1601
他のコマンドの処理
1611
Yes
ヒット?
1612
No
キャッシュメモリ上のヒットしたデータをホストへ転送
1602
データ転送指示
No
ハーフヒット?
1613
Yes
1614
新しセグメントを生成
1615
ハーフヒットしたセグメントを更新
先読み指示
最大先読み量をディスクフォーマッタに設定
1616
先読み
1603
No
ホスト要求データが用意できた?
1604
1617
セグメント情報を仮更新
セグメント更新割り込み
Yes
先読み
1605
No
最大先読み量に先読み量が達した?
1606
Yes
先読み停止
1607
先読み停止割り込み
1618
セグメント情報を更新
終了

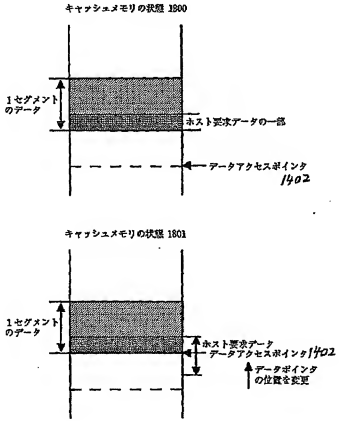
【図１７】

図17


キャッシュメモリの状態 1700
1セグメントのデータ
ホスト要求データの一部
データアクセスポインタ 1402
キャッシュメモリの状態 1701
データアクセスポインタ 1402
ホスト要求データ
1セグメントのデータ
データポインタの位置を変更

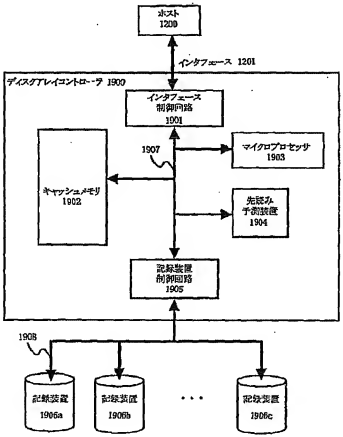
【図１８】

図18


キャッシュメモリの状態 1800
1セグメントのデータ
ホスト要求データの一部
データアクセスポインタ
1902
キャッシュメモリの状態 1801
1セグメントのデータ
ホスト要求データ
データアクセスポインタ1902
データポインタの位置を変更

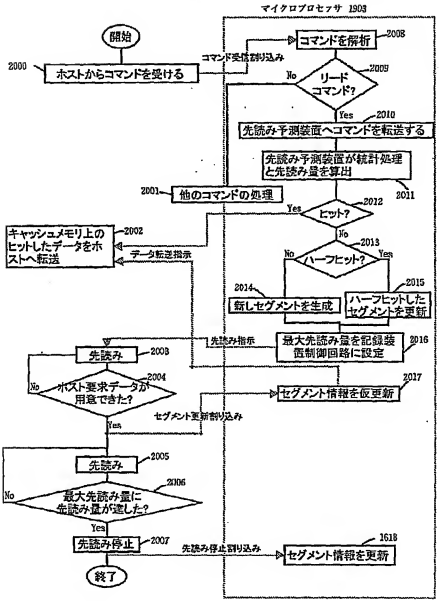
【図１９】

図19


ホスト
1200
インタフェース 1201
ディスクアレイコントローラ 1900
インタフェース
制御回路
1901
1907
マイクロプロセッサ
1903
キャッシュメモリ
1902
先読み
予測装置
1904
記録装置
制御回路
1905
1908
記録装置
1906a
記録装置
1906b
記録装置
1906c

【図２０】

図20



フロントページの続き

(72)発明者　本田　聖志
神奈川県川崎市麻生区王禅寺1099番地　株式会社日立製作所システム開発研究所内

(72)発明者　市川　正敏
神奈川県川崎市麻生区王禅寺1099番地　株式会社日立製作所システム開発研究所内

(72)発明者　高安　厚志
神奈川県小田原市国府津2880番地　株式会社日立製作所ストレージシステム事業部内

(72)発明者　酉川　学
神奈川県小田原市国府津2880番地　株式会社日立製作所ストレージシステム事業部内

Ｆターム(参考) 5B005 JJ13 MM11 NN22 VV03
5B065 BA01 CE12 CH05
5D044 AB01 BC01 CC04 FG10 FG30
NH02 HL02